



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/513,058	02/24/2000	Anita Wai-Ling Huang	AM999074	5519
23334	7590	04/06/2006	EXAMINER	
FLEIT, KAIN, GIBBONS, GUTMAN, BONGINI & BIANCO P.L. ONE BOCA COMMERCE CENTER 551 NORTHWEST 77TH STREET, SUITE 111 BOCA RATON, FL 33487			BASEHOAR, ADAM L	
			ART UNIT	PAPER NUMBER
			2178	
DATE MAILED: 04/06/2006				

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	Application No.	Applicant(s)	
	09/513,058	HUANG ET AL.	
	Examiner	Art Unit	
	Adam L. Basehoar	2178	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

1) Responsive to communication(s) filed on 23 January 2006.  
 2a) This action is FINAL.                    2b) This action is non-final.  
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

4) Claim(s) 1,2,5-12,14,17,18,20,21,25 and 26 is/are pending in the application.  
 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.  
 5) Claim(s) \_\_\_\_\_ is/are allowed.  
 6) Claim(s) 1,2,5,6, 10-12,14,17,18,20,21,25 and 26 is/are rejected.  
 7) Claim(s) 7-9 is/are objected to.  
 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

9) The specification is objected to by the Examiner.  
 10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).  
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
 a) All    b) Some \* c) None of:  
 1. Certified copies of the priority documents have been received.  
 2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

1) <input type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. _____
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.	5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)
	6) <input type="checkbox"/> Other: _____.

## **DETAILED ACTION**

1. This action is responsive to communications: The Appeal Brief Filed 01/23/06.
2. The rejection of claims 7-9 under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99) have been withdrawn as necessitated by Appeal Brief.
3. The rejection of claims 1-2, 5-6, 9, 12, 14, 17-18, 20-21 remain under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99).
4. The rejection of claims 10-11 and 25-26 remain under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99) in view of Microsoft Press Computer Dictionary, Microsoft Press, 1997, pp. 309.
5. Claims 1, 2, 5-12, 14, 17, 18, 20, 21, 25, and 26 are pending in the case. Claims 1, 10, 12, 17, and 25 are independent claims.

### ***Claim Rejections - 35 USC § 101***

6. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

7. Claims 1, 10, 12, 17, and 25 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Independent claims 1, 10, 12, 17, and 25 are considered non-statutory subject matter because they do not produce a tangible result. The claims are merely directed to determining the equivalence of two documents. Looking at claim 1, an example of a tangible result that would partially overcome the rejection resides in dependent claim 6, wherein as a result of determining equivalence the second document metadata summary was removed. However this result would only cure the determining that the

two documents are equivalent. Either dependent claim 7 or 8 would be needed to show a tangible result wherein the documents are determined to be distinct (i.e. storing a zero value in the created metadata table). Appropriate correction is required.

***Claim Rejections - 35 USC § 103***

8. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

9. Claims 1-2, 5-6, 12, 14, 17-18, 20-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99).

-In regard to substantially similar independent claims 1, 12, and 17, Brown et al teach a method, system, and computer program product comprising:

receiving a first and a second document (column 4, lines 59-66)(Fig. 3B);

a metadata parser generating a first and second metadata summary for each of the received documents (column 7, lines 40-67)(Fig. 3B), wherein metadata summaries included a plurality of sub-trees(i.e. Intrinsic and Non-Intrinsic Attributes)(column 7, lines 48-67)(Fig. 3B: 370 & 380) wherein each sub-tree could include a plurality of nodes (i.e. Intrinsic (relevance score, title, size) and Non-Intrinsic (location which includes filename (Fig. 2b));

storing the metadata summaries in a repository (Fig. 1 & Fig. 4: 410);

a summary consolidator for comparing the metadata summaries on a structural level by comparing the corresponding sub-tees (Fig. 4: 420, 430 & 440, 450);

identifying the first and second documents as distinct if the structure of the sub-trees are not equivalent (Fig. 4: 455);

if the structures are equivalent performing the steps of:

comparing the first and second metadata summaries on a textual level (i.e. by comparing relevance score (Fig. 3B: 375) which was calculated by information retrieval algorithm as a function of the query (Fig. 3A) and the contents of the document (column 7, lines 55-67)) by comparing the textual content from the first and second document contained in the metadata summaries (Fig. 4: 420); and

identifying the first and second documents as distinct if the textual content within the sub-trees are not equivalent (Fig. 4: 455).

Brown et al do not specifically teach comparing the metadata summaries on a structural level before comparing the textual content (e.g. “relevance score” “title”). It would have been obvious to one of ordinary skill in the art at the time of the invention for Brown et al to have compared the sub-tree structures of the metadata summaries of the documents to determine document distinctness before comparing the textual content of the documents to determine distinctness, because by checking the structure of sub-trees first, it could more quickly be established with the minimum amount of comparisons that the compared documents were distinct from one another (e.g. Referring to Fig. 3B it is shown that there are fewer sub-tree elements than there are textual values within the elements. Thus it was obvious if one document maintained a “Title” structure and a compared document did not then those two documents would be distinct and said distinction would be obtained more quickly than parsing and

comparing the one document title of “HEAVY RAINS FLOOD FARMS IN MID-WEST” with no title of another document)

-As per dependent claims 2, 14, and 18, Brown et al further disclose comparing the first and second metadata summaries on an attribute level (author, creation date, length, size, location, abstract, etc (columns 1 and 6, lines 64-67 and 51-54)) by comparing attribute values within the sub-trees of the first and second metadata summaries and identifying the first and second documents as distinct if the attribute values within the sub-trees are not equivalent. (Fig. 4: 455). Similar to the independent claims discussed above, Brown et al do not specifically teach wherein the attribute comparison was made after the structural comparison and before the textual comparison. Similarly, it would have been obvious to one of ordinary skill in the art at the time of the invention for Brown et al to have compared the attributes of the metadata summaries after the structural comparison and before the textual comparison, because an attribute comparison provided the second quickest determination of distinctness between the documents (i.e. checking the attribute values stored in the metadata summaries required more processing than checking a few nodes in the metadata summaries and required less than checking the textual content of the document metadata summaries).

-As per dependent claims 5 and 20, Brown et al further disclose noting the current pair of hit-list documents as duplicates if the text content (column 7, lines 55-67) (Fig. 3B: 375) are equal (Fig. 4: 432) (i.e. the structure and non-text attributes have been compared but alone can not determine duplicates)(Fig. 4: 460).

-As per dependent claims 6 and 21, Brown et al further disclose one embodiment of *deleting duplicates* from the hit list group (column 8, lines 20-22: Fig. 8).

10. Claims 10-11 and 25-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99) in view of Microsoft Press Computer Dictionary, Microsoft Press, 1997, pp. 309.

-As per independent claims 10 and 25, Brown et al teach the method of claim 1 as stated above, and further disclose receiving not just two, but *two or more entries* (column 3, lines 54-55) of documents and generating metadata hit-list summaries for each received document (Fig. 3b). Brown et al do not teach that when receiving a plurality of documents to group the metadata hit-list summaries according to the mime-type designation of each document. The Microsoft Press Computer Dictionary teaches us that MIME-types describe the contents of a document which can be used to interpret the content of a file over the internet (pp. 309). It would have been obvious to one of ordinary skill in the art, to have used Brown et al method for identifying duplicate documents from search results without comparing document content and grouping the hit-list summaries by MIME-type, because documents of different MIME-types wouldn't have to be compared as their intrinsic attributes would be wholly different, i.e. a regular text/plain MIME-type couldn't be a duplicate of a text/html MIME-type, and thus this would increase efficiency by significantly reducing the number of hit-list summary comparisons. Note the above rejections of independent claims 1, 12, and 17 with regards of comparing the metadata summaries on a structural level to determine document equivalence.

-As per dependent claims 11 and 26, Brown et al teach generating metadata summaries for each received document (Fig. 3B). Brown et al and the Microsoft Press Computer Dictionary do not teach grouping more subsets of hit-list summaries by MIME-type. As it was taught above in the rejection of independent claims 10 and 25, it would have been obvious to have grouped as many MIME-type subsets as were required by the plurality of documents for the purpose of increased efficiency by significantly reducing the number of hit-list summary comparisons across MIME-types.

***Allowable Subject Matter***

11. Claims 7-9 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

***Response to Arguments***

12. Applicant's arguments filed 01/23/06 have been fully considered but they are not persuasive.

-As noted above and upon further consideration of the claimed subject matter in view of the referenced prior art, dependent claims 7-9 are considered allowable subject matter.

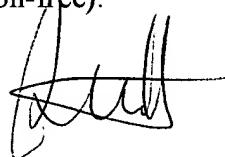
-Please note the Examiner's previous response to arguments with regards to independent claims 1 and 10, as Applicant appears to be stating similar arguments to those presented in earlier responses.

***Conclusion***

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Adam L. Basehoar whose telephone number is (571)-272-4121. The examiner can normally be reached on M-F: 7:00am - 4:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Steve Hong can be reached on (571) 272-4124. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



ALB

STEPHEN HONG  
SUPERVISORY PATENT EXAMINER